

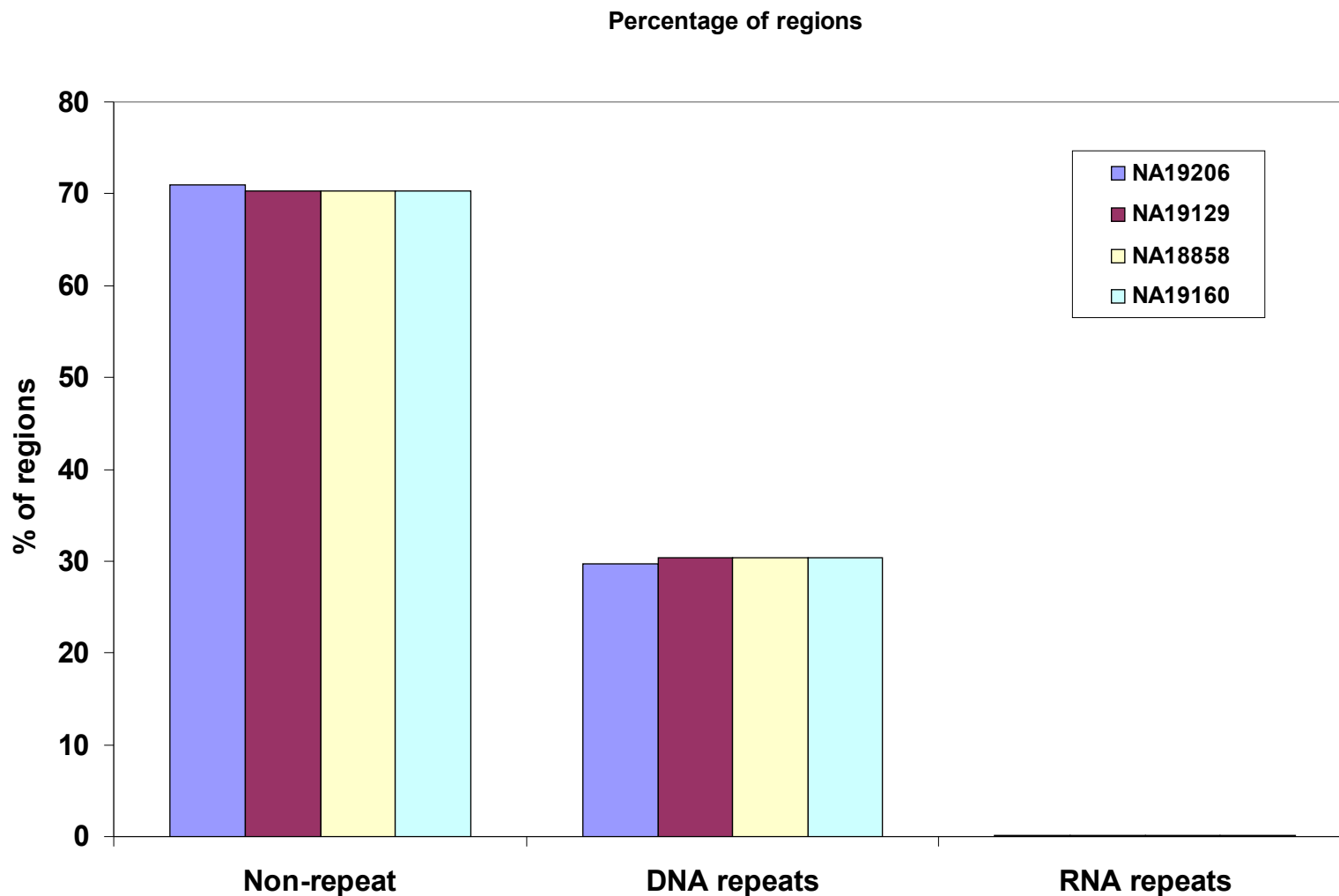
Repetitive DNA sequences and their effects on gene expression

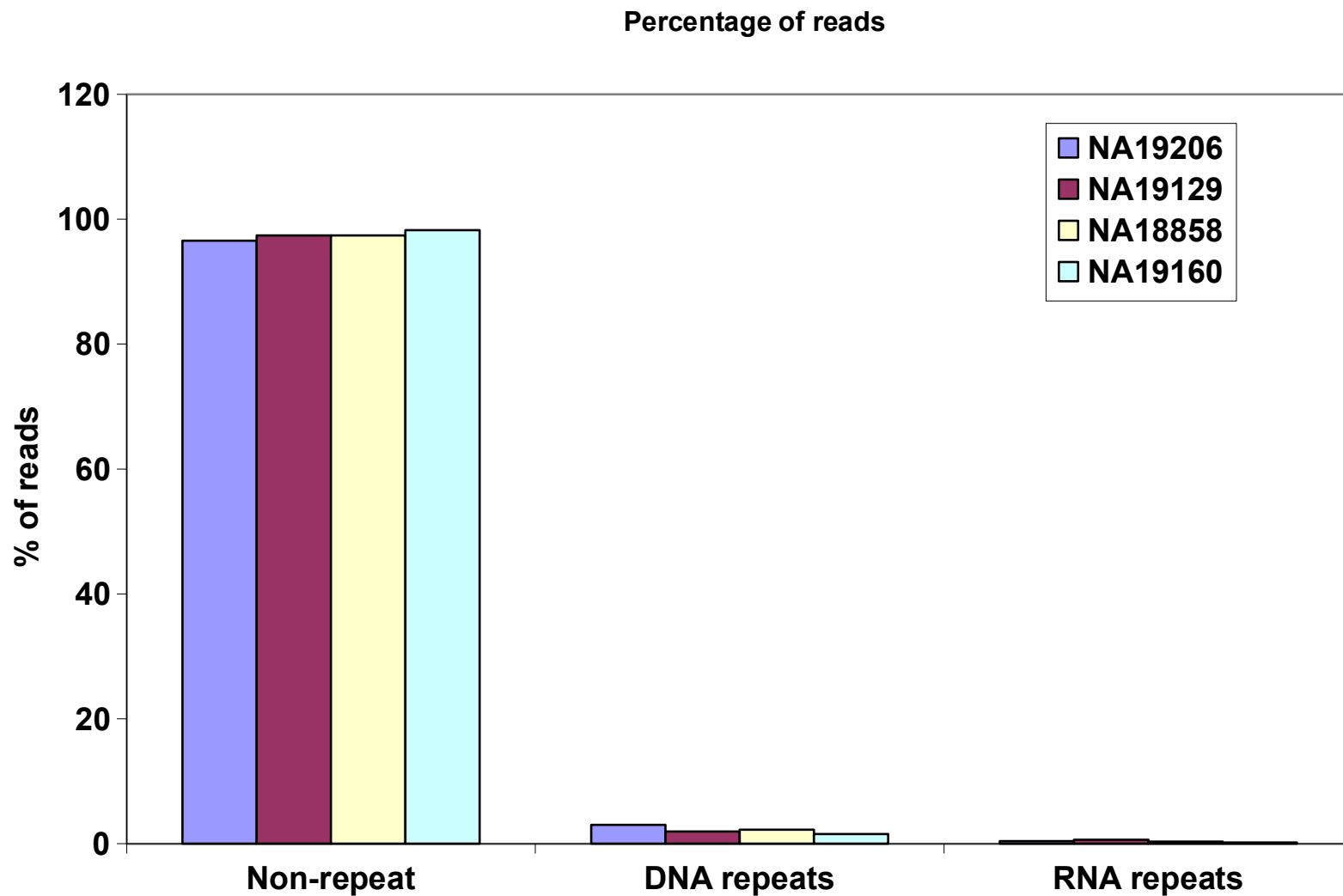
**Irina Pulyakhina,
Department of Human Genetics
Leiden University Medical Center
Leiden, the Netherlands**

- So far only unique regions in the genome are studied
- Repetitive regions represent large part of the genome and are probably expressed

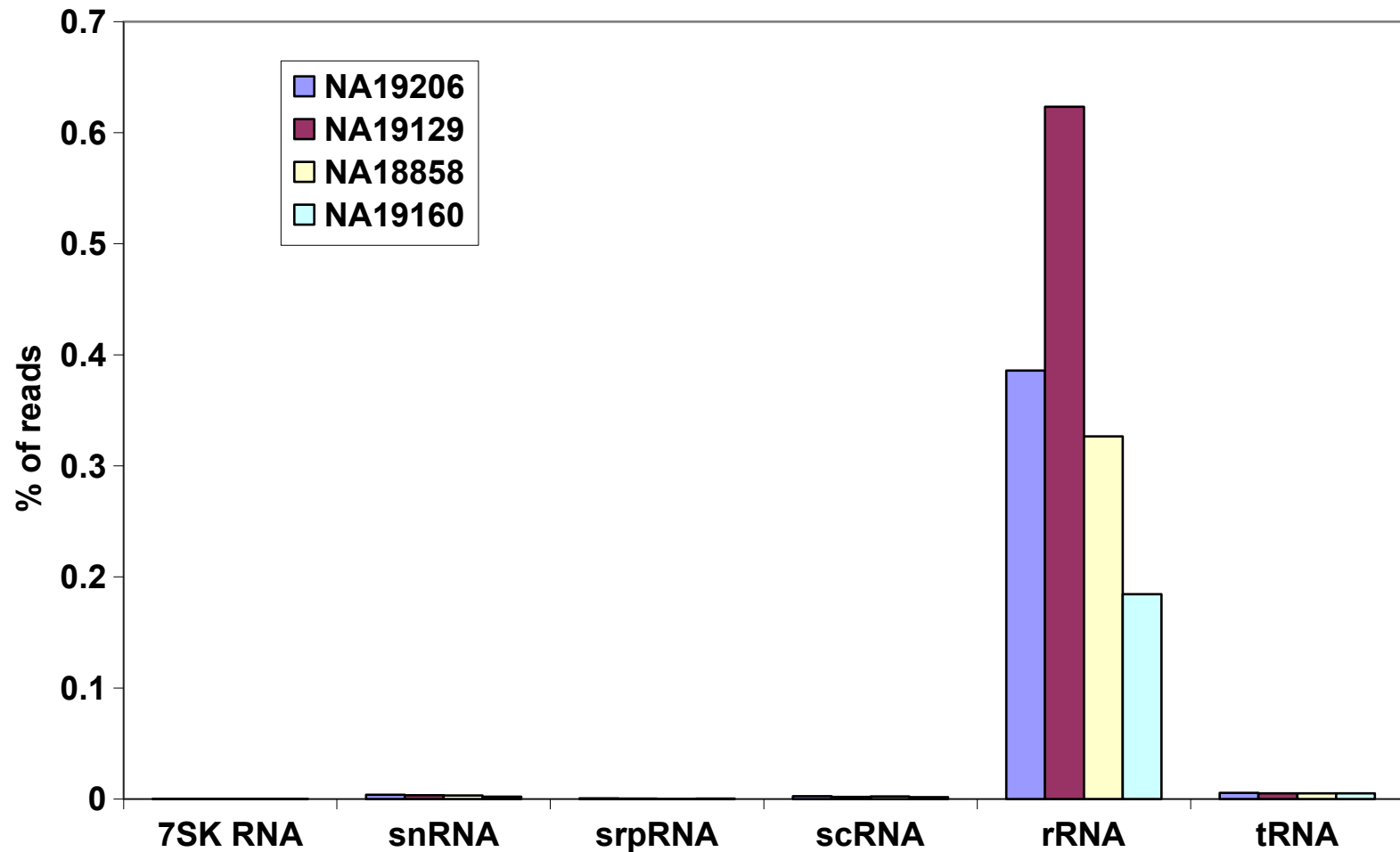
- Defined as regions of contiguously covered nucleotides in the genome
- Based on GEM alignment
Evaluated 4 samples from 4 different centers
- Repeat regions taken from UCSC Repeat Masker track

Expressed regions overlapping with repeats

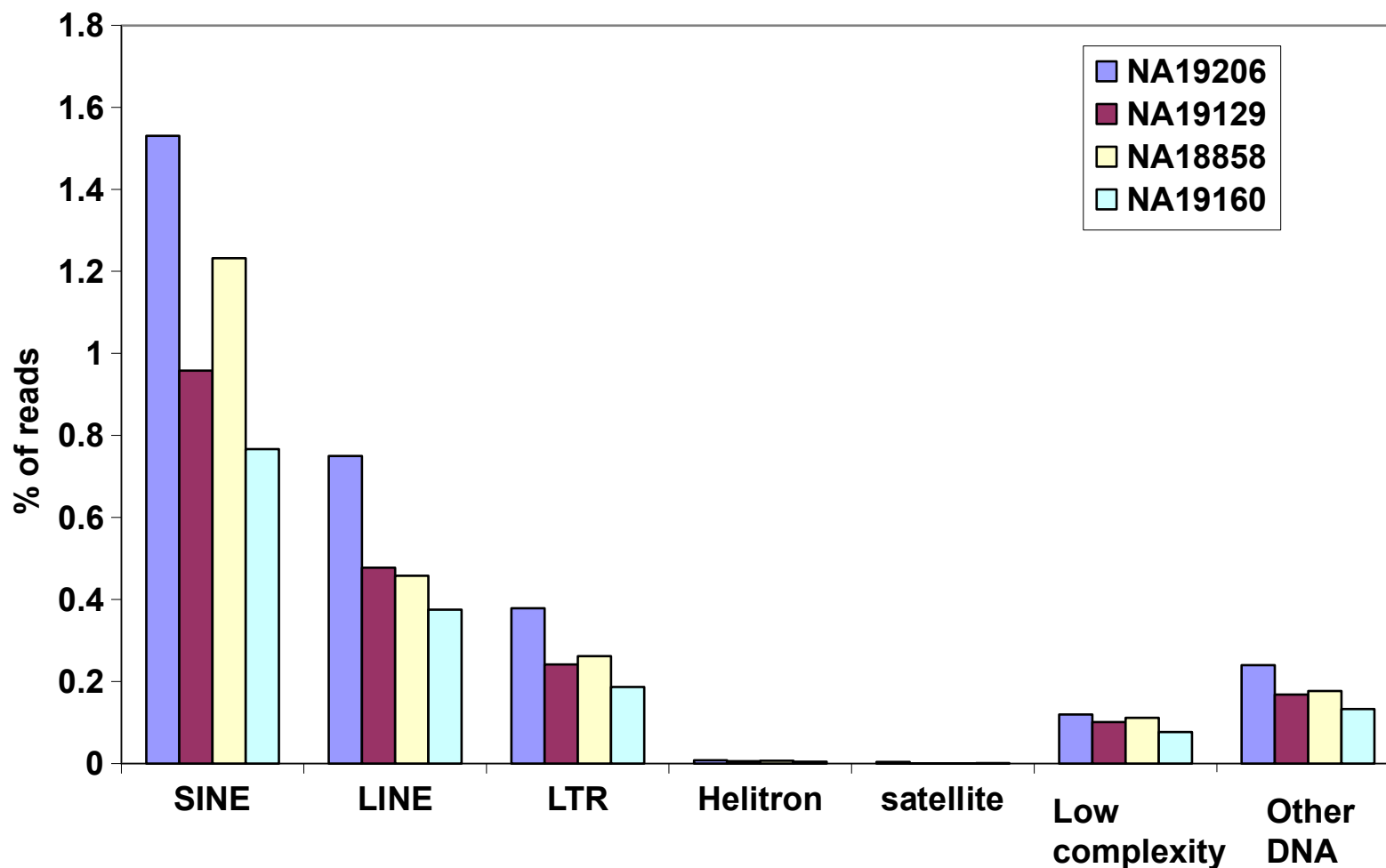




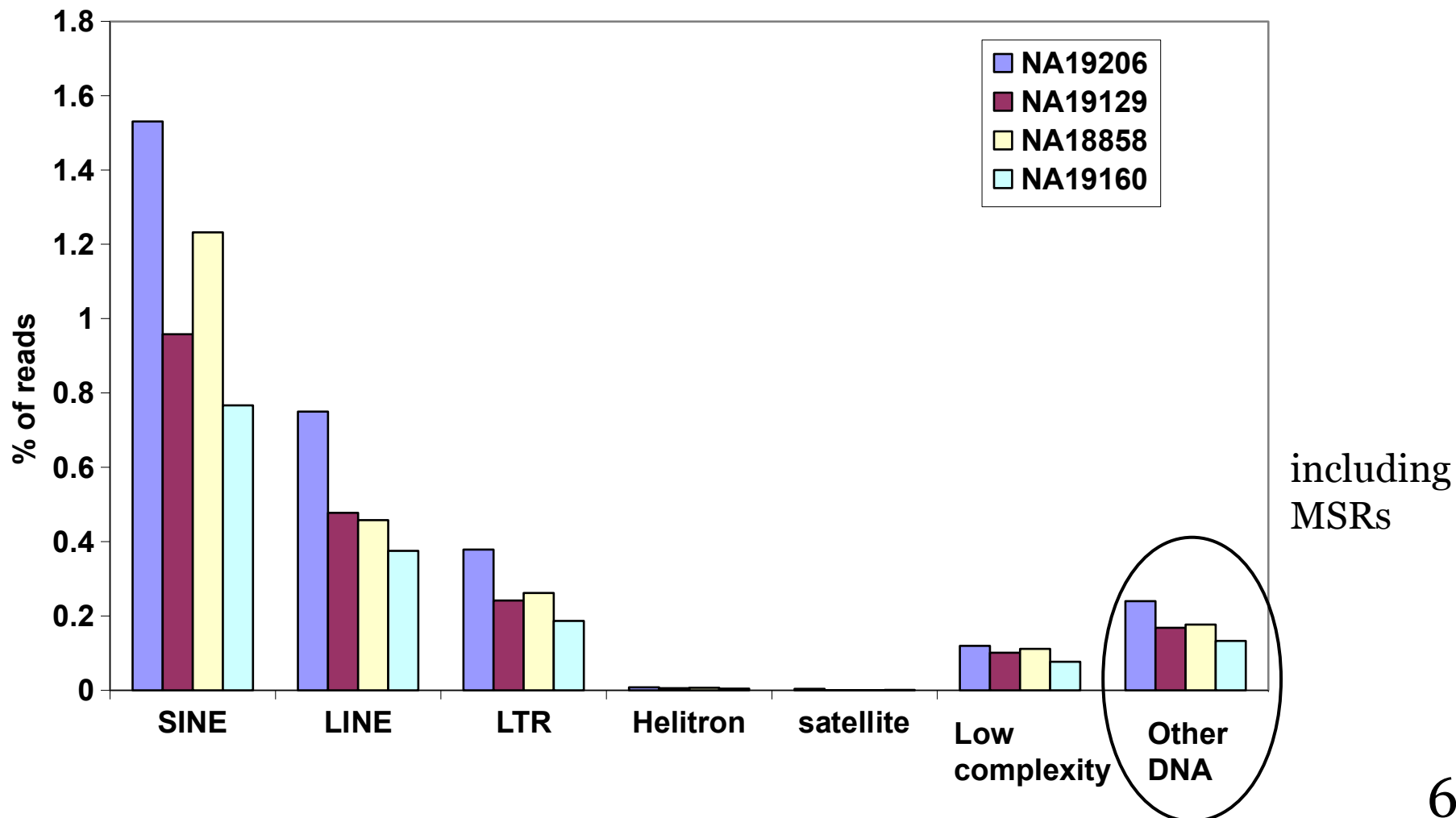
Percentage of reads mapping to RNA repeats



Percentage of reads mapping to DNA repeats



Percentage of reads mapping to DNA repeats



Repetitive DNA large part genome.

Tandemly repeated sequences:

Microsatellites (1-7 bp)

Minisatellites (<100 bp)

Macrosatellites (>100 bp)

Repeat (array) size often hundreds of kilobases.



Structure and function not well studied yet.

Best studied MSR is D4Z4 (<11 units FSHD).

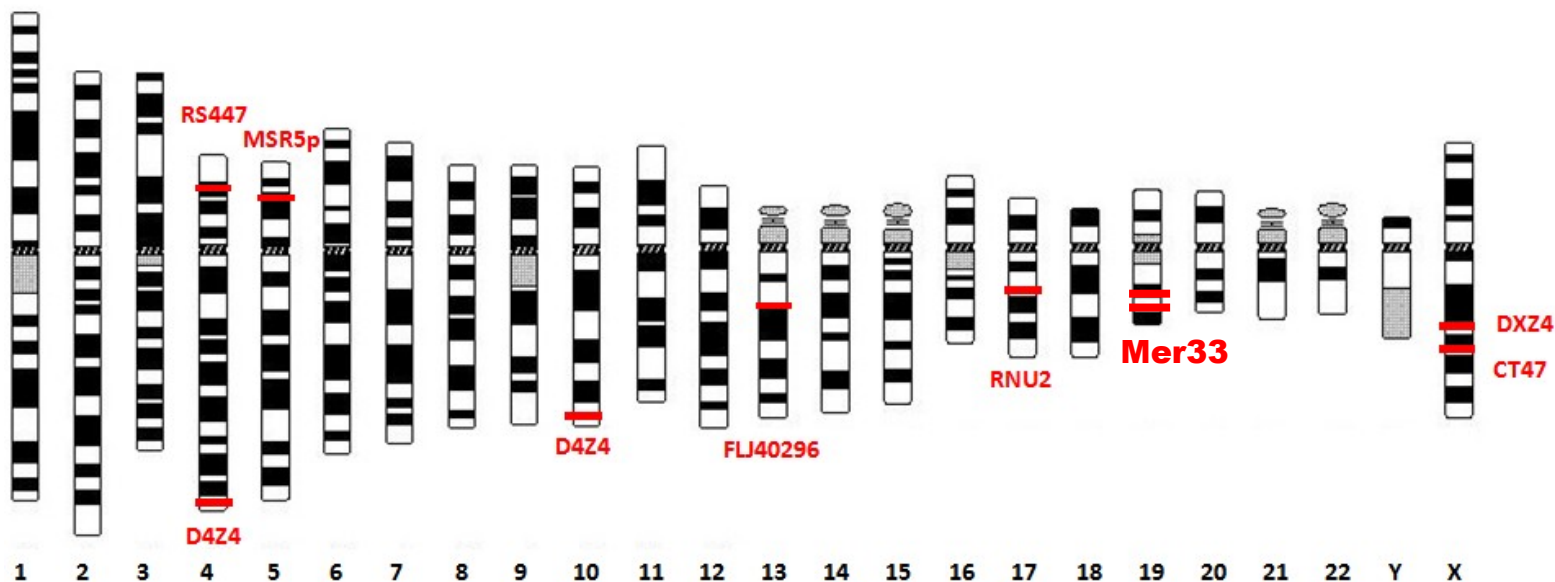
DXZ4 suggested role in X inactivation (Chadwick et al. 2008).

Structure – technical limitations.

10 MSRs on different chromosomes

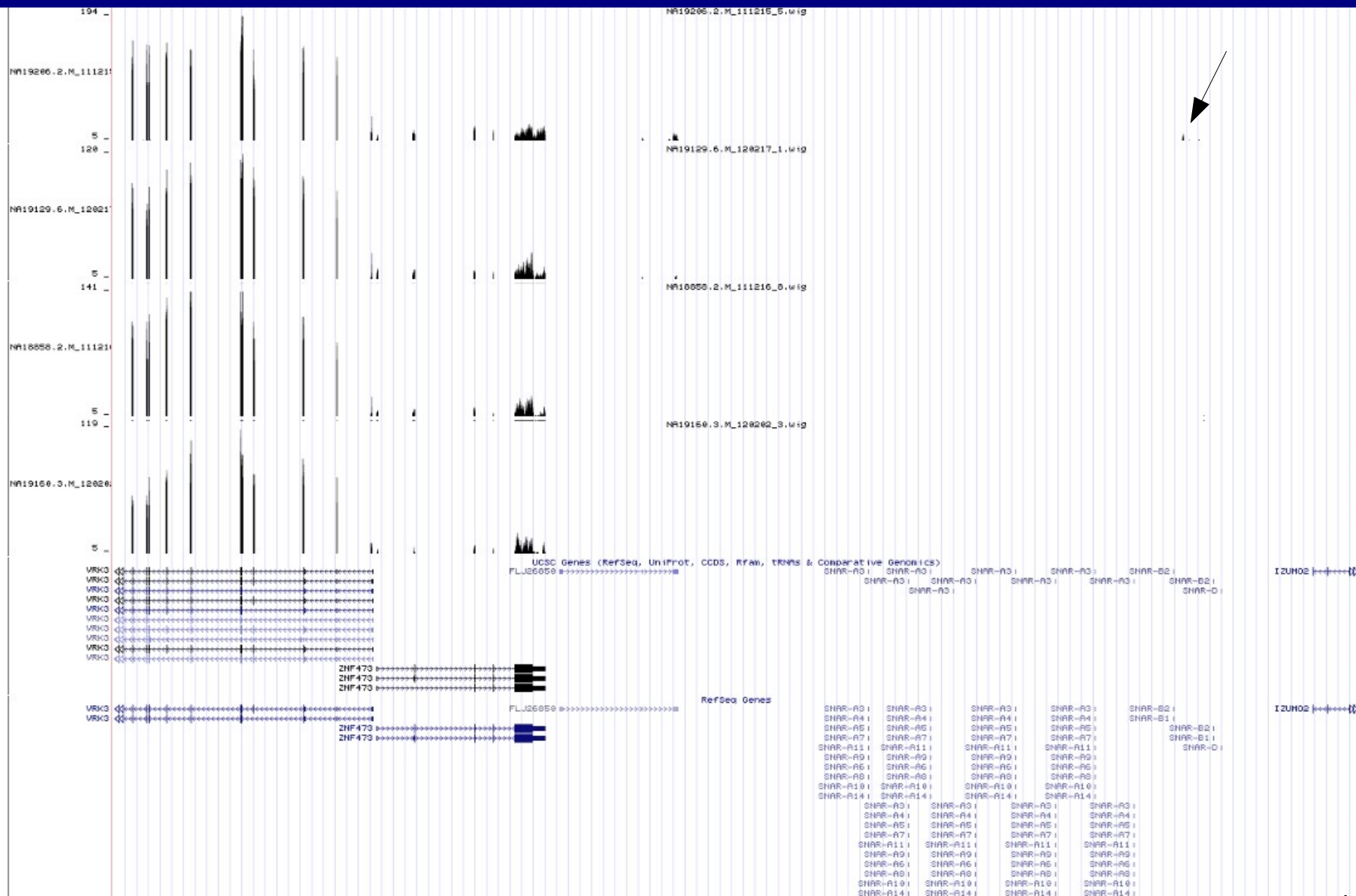
8 autosomal, 2 X-chromosomal

3.0 kb – 6.6 kb



Name	Locus	Unit (kb)	Literature	This study	Minimum (units)	Mitotic (%)	Meiotic (%)
RS447	4p16.1	4.7	20-103	11-113	11	0.4	-
D4Z4	4q35.2	3.3	1-113	6-113	1	-	-
MSR5p	5p15.1	3.4	10-98	8-131	8	0.4	-
D4Z4-like	10q26.3	3.3	1-105	2-105	1	-	-
FLJ40296	13q21.1	6.6	5-20	3-30	3	0.7	1.7
RNU2	17q21	6.1	5-41	5-63	5	1.5	-
DXZ4	Xq23	3.0	12-100	18-120	12	1.1	-
CT47	Xq24	4.8	13	4-17	4	0.7	-

- Almost absent on mRNA, except for DXZ4
- **miRNA???**

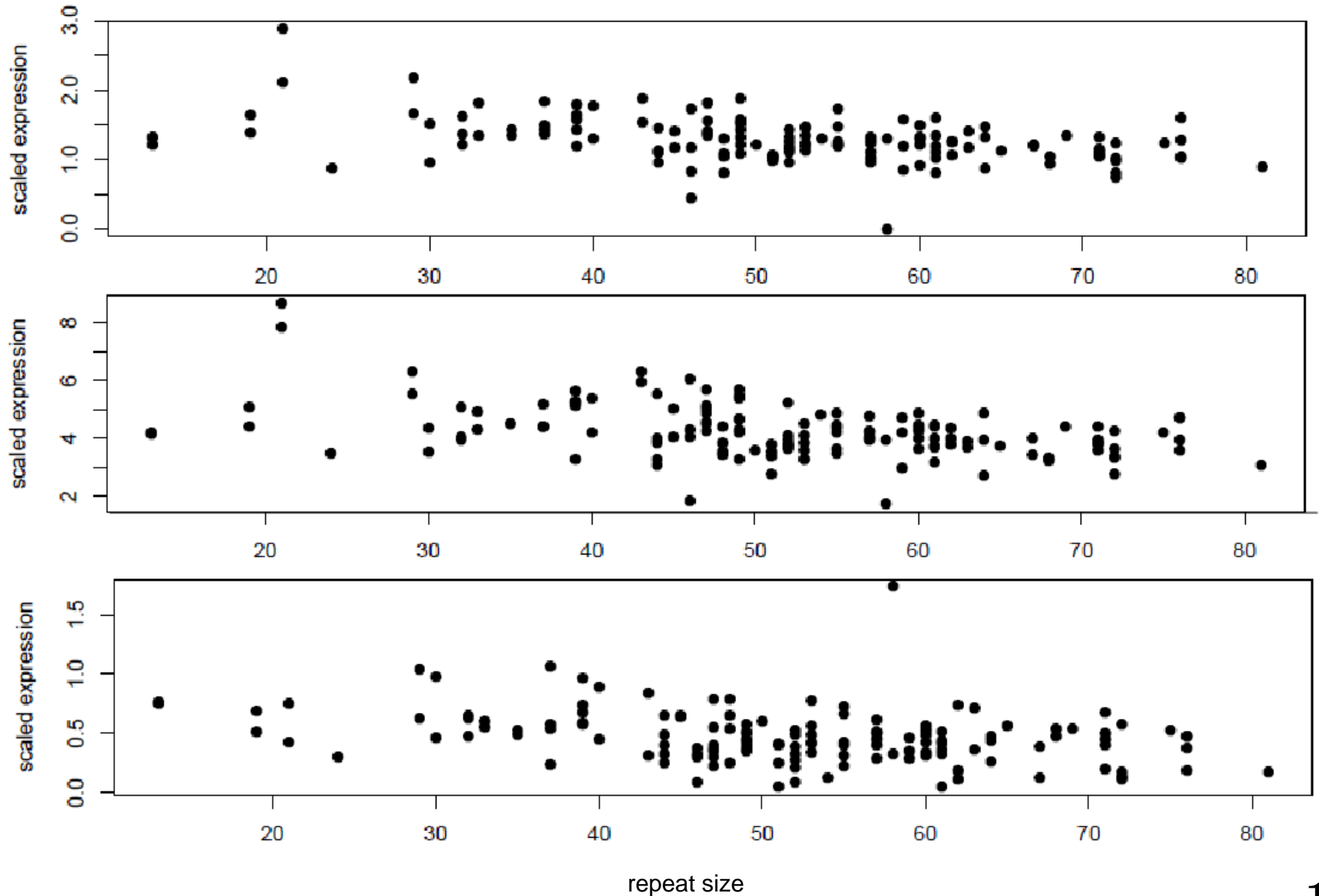


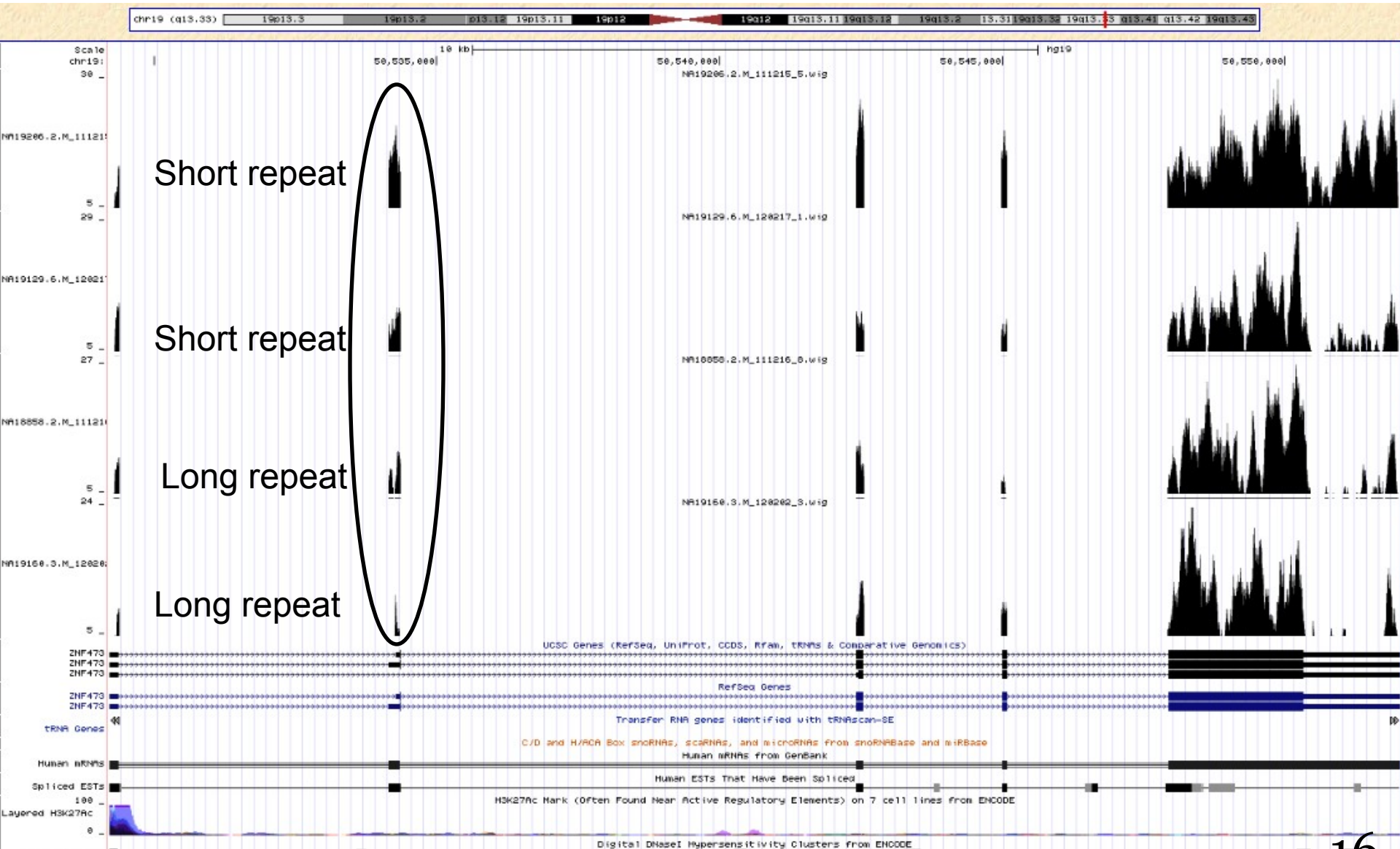
- Exon level counts GEM alignment
- Scaled by library size
- Selection of exons within 2Mb of repeat
- Linear model per exon
 - Expression $\sim \alpha + \beta * nrepeats$
 - $\text{Sqrt}(\text{expression}) \sim \alpha + \beta * nrepeats$

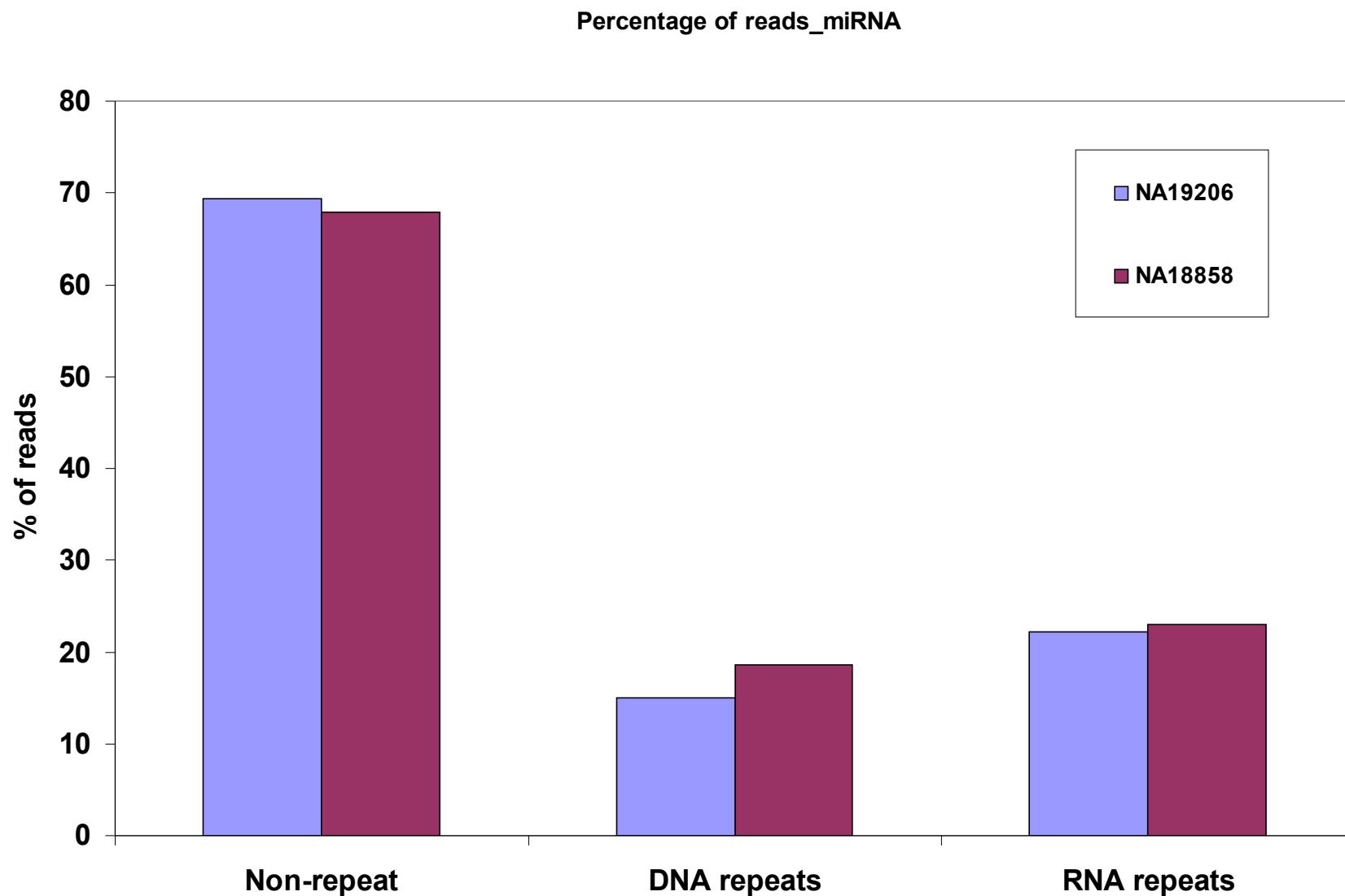
- 3 exons with adjpvalue < 0.05
 - sqrt or linear

- All three are very close to the repeats, just upstream.

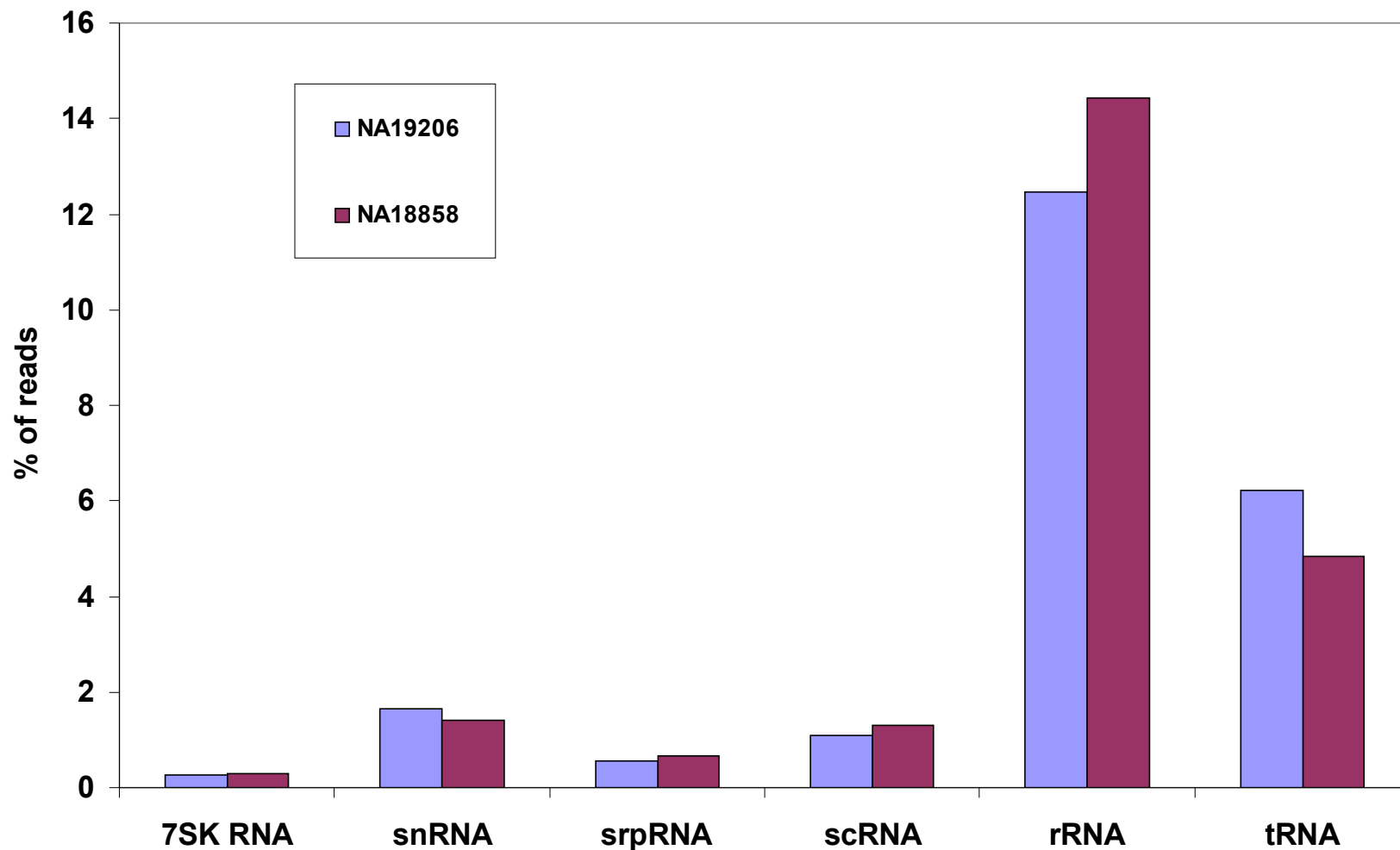
Correlation between exon expression and repeat size



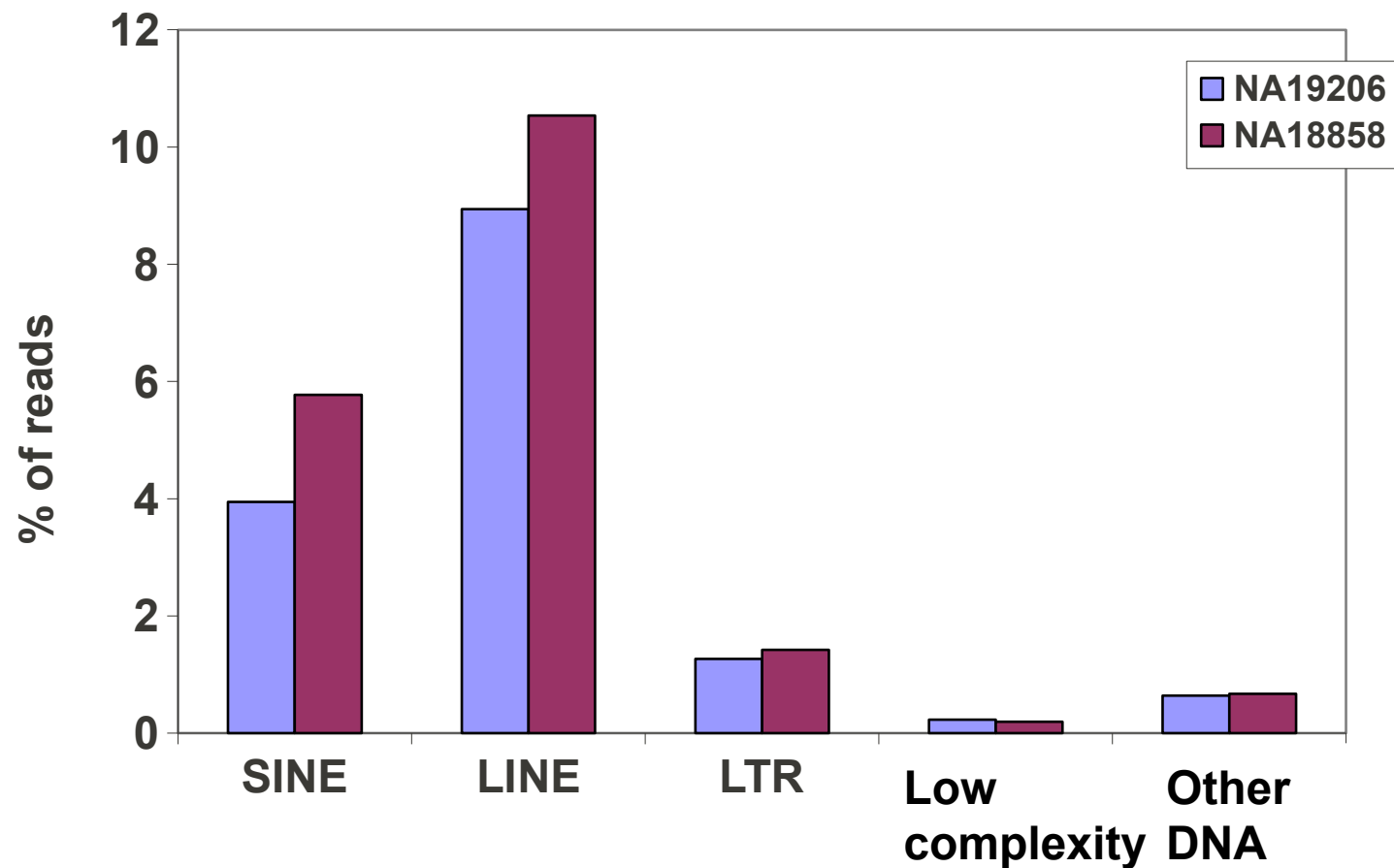




Percentage of reads mapping to RNA repeats



Percentage of reads mapping to DNA repeats



- 1. Considerable expression of DNA repeats.**
- 2. These repeats can be expressed themselves or influence expression/alternative splicing of neighboring genes.**
- 3. The effect single nucleotide variants and structural variants on the expression of small RNAs warrants further investigation.**

- 1. Study the influence of number of repeat units on expression for other MSRs.**
- 2. Validation and mechanistic studies of the effects of repeat unit length on gene expression.**
- 3. Study the effect of SNV / SVs on expression of structural and repetitive RNAs. Consider different alignment strategies dealing with repetitive RNAs, and k-mer counting.**

**Peter 't Hoen
Mireille Schaap
Judith Balog
Silvere van der Maarel
Henk Buermans**